DATA PROCESSING DIGEST

1140 South Robertson Blvd., Los Angeles 35, California

a publication of Canning, Sisson and Associates, Inc.

VOLUME 6 NUMBER 6

JUNE, 1960

General Information

HOW MUCH DATA SCREENING TO MINIMIZE COMPUTER-DETECTED ERRORS

Phillips S. Cruttenden, U.S. Air Force, Cleveland Air Procurement District
N.A.A. BULLETIN, April 1960; pages 85-89

"How much preliminary screening and auditing work should be done before transactions are entered on a computer?... Theoretically, the computer should be programmed to take practically all transactions in random order. It should detect most of the errors and omissions and make necessary corrections, if possible, without the necessity of rejecting the transaction and printing out the error on an error log." Since this ideal is seldom achievable, some preliminary screening is likely to be necessary. If the error rate should increase, the backlog of unprocessed work also increases, and it takes more time to process corrections than it does to process a transaction initially, to say nothing of the cost of computer re-runs. Accurate preparation and screening of source documents will make the computer error rate approach zero, and the cost of error correction will be low. However, if preliminary screening is eliminated, the number and cost of errors detected by the computer could rise to unacceptable heights.

To find the break-even point between the preliminary screening costs and computer-detected error correction costs, reliable data must be collected. This includes data on the frequency of errors developed by 1) finding the cost of performing preliminary screening (A), 2) determining the frequency of errors found in the screening process (B), 3) finding the cost of doing the error correction work (C), 4) finding the frequency of error detection by the computer minus the errors which preliminary screening clerks did not attempt to find, check or verify (D), 5) decide how much error reduction is desired (E), and 6) include an unknown (K) which is the amount of cost that could be added to preliminary screening while saving an equal amount of error correction costs at the going rate for errors.

We now have the following:

A--Preliminary screening costs
B--Preliminary error rate
C--Error correction costs
D--Computer error rate
E--Change in error rate
K--Break-even costs

CONTENTS

The relationship of these factors is:

(A+K) is to (B+E) as (C-K) is to (D-E)

From this, we develop the equation for finding the value of K, which is:

$$K = A(E-D) + C(B+E) \div (B+D)$$

Four ways to reduce errors

To reduce errors: 1) step up the training of source document clerks, or 2) revise computer programs to correct more errors automatically, or 3) stimulate screening clerks to detect more errors, or 4) add more people to the screening unit or add overtime. How much can these costs be increased in relation to the desired reduction in computer error rate? By substituting the values in the above equation, a break-even cost may be found which can be added to the cost of the preliminary screening work while saving an equal amount of error correction costs.

"How much can we increase these costs and still obtain a .02% reduction in the computer error rate with a comparable saving? We have accumulated the following costs and error frequencies:

A--Preliminary screening costs	$420.00
B--Preliminary error rate	.06%
C--Error correction costs	765.00
D--Computer error rate	.11%
E--The amount of error reduction proposed	.02%

Where is the break-even point?

By substituting these values in the equation, we obtain break-even costs which can theoretically be added to the cost of the preliminary screening work while saving an equal amount of error correction costs.

$$K=420(.02-.11)+765(.06+.02)\div(.06+.11)=\$137.65$$

Another use of the break-even formula is to find the amount of change in the computer error rate when the amount of additional preliminary screening costs is known. The equation to find E when K is known is derived from the same basic equation and becomes:

$$E=D(A+K)-B(C-K) \div (A+C)$$

"In practice, it is best not to spend much money on preliminary screening and checking of data that the computer has been programmed to check. The computer can do the work much more completely and accurately."

ERROR CONTROL IN DATA INPUT SYSTEM DESIGN—PART I

Roger L. Sisson, Canning, Sisson and Associates, Los Angeles
AUTOMATION, DECEMBER 1958; pages 40-45

Data processing systems involve a series of steps, each of which is complex and must be executed properly, if the end results--management reports--are to be worthwhile. Data recording and data communications have had less improvement than the processing, storing and presenting of data. However, significant improvements can be made in many data input systems, and many small businesses can obtain direct savings and improved reporting through a better input system, even though they cannot justify EDP.

Control errors at the recording point

Data originally recorded by humans (the most prevalent system) contains an error rate of from one to five percent. As a general rule, wherever there are employees, such as timekeepers, who devote a large part of their time to data recording, there is a good chance that a mechanized system will pay off. A procedure is given for analyzing the problem and for designing a system which will control errors and reduce the time between recording and processing economically, based on the example of data processing in a department store. Steps in analyzing an existing system and determining the basic requirements are:

1. List all of the types of transactions that are recorded.

2. Determine by area the volume of transactions--both average and peak.

3. Obtain a list of the fields, or items of information required for each type of transaction. For each field determine: the number of characters required (average and peak), and what per cent of the time the field is required.

Analyze the input environment

In studying the inputs, include inputs which serve auditing functions, such as use of the employee's signature on a job report as proof of his working on the job. A simple way of gathering information about present inputs is to collect all of the forms presently used, both a blank copy, and an example filled out with a complex case. Include important telephone inputs, also. Steps 1, 2, and 3 should be repeated for the outputs of the data input system where the output forms are different from the input forms.

4. List all of the types of errors made in the input, and for each, the number of that type of error per transaction.

5. Observe the environment in which the data is recorded (dust, noise, vibration, etc.) and whether the process is manual or mechanical.

6. Consider the psychological factors: .organization position of the people who must record the data, the attitude of the operators toward the clerical aspects of their job, amount of training in clerical procedures the workers have had.

7. Survey all possible communications methods, especially inter-city or inter-plant. For each possible communication system obtain: cost per character or message, characters per second, and error rate under normal operating conditions.

The design of the input system will be influenced largely by the design of the error control system. This is based on the principle of redundance, that is, upon sending more information than would be required in an error-free system. The redundant information must be added in a controlled way, however. Each preassigned field can carry a precalculated checking digit. Groups of fields can be checked by batch, control, or hash totals. "A broader look at the data flow will often permit the comparison of data from two inputs or the input and a file to develop consistency and reasonableness checks." The following design problems must be faced:

How to figure cost of errors

1. Determine where error checks should be made. "From one point of view the check should be made as early as possible in the series of steps starting with recording of data and continuing to input and processing. This permits the correction loop. . . to be as simple and inexpensive as possible. . . . On the other hand, if the error rate is low, it may be better to wait until the data get to a computer to detect errors since the cost of handling the non-error cases will then be lower."

2. As to how much checking is advisable, the general rule is that the cost of error detection and correction should not be greater than the cost of letting the error get through to affect the files and output reports.

The cost of errors may be figured by the following procedure: "Design several possible systems, both manual and mechanized. Study each to determine how the operating cost of the system. . . will vary with varying amounts of error detection and correction. It should then be possible to plot a curve. . . which shows how the cost of error control will vary with the error rate. From this curve, management can pick an acceptable operating point."

THE SECOND PHASE OF COMPUTER PROBLEMS

R. Lee Paulson, Esso Standard Oil Co., Baton Rouge, Louisiana
JOURNAL OF MACHINE ACCOUNTING, April 1960; pages 14-21

The computer application cycle tends to repeat itself in the following pattern:

<u>Phase I</u>
1. Idea and logic study
2. Programing and testing
3. Parallel (initial) production
4. Non-routine production

<u>Phase II</u>
5. Routine production
6. Re-evaluation
7. Proposed major change in concept.

The fifth step can continue for a long period of time before it reaches the re-evaluation step in the cycle. Most of the applications will be at various stages in the cycle. When the bulk of the original applications designed for the computer reach the second phase of the cycle, the EDP installation will find itself with new and different problems that must be defined and resolved. These are:

1. Increasing utilization of time. Requests for machine applications do not diminish, but continue steadily. To maintain his routine production basis the manager may: a) secure more machine capacity by buying or renting an additional computer; b) purchase the needed machine time from a service bureau; c) squeeze more capacity from his own machine. The latter will almost always be more economical.

Better scheduling adds capacity

To get needed capacity, he must begin by establishing priorities. All applications, old as well as new, must be constantly reviewed and re-evaluated. All marginal runs should be backed out of the EDPM system to a lower cost method. Punched card equipment can take considerable pressure off the computer. Through efficient programing, the manager can do much to increase his computer's scheduled time potential. Good maintenance programing, methods work, flexible scheduling, up-to-date equipment, are all ways to gain time. Expansion to two or three shifts should be a last resort.

2. Realizing maximum benefit from routine production. Maintenance programing can make a significant contribution to EDPM efficiency. If the machine application is not kept up to date, the output becomes less and less useful because it is out of touch with actual conditions.

Another area in which progress must be made is that of systems integration. Programing personnel must realize the potentials of higher levels of systems refinement.

3. Increasing need for cooperation and coordination. One of the vital parts of any production run is the feeding of source data to the machine. When the EDPM system becomes sophisticated to the extent that the output from one production run provides the raw data for another run which in turn feeds additional runs, the timing of feed becomes critical. Late or improper feed for any particular run may create a chain reaction affecting a whole series of runs and projects. Close and active coordination becomes vital.

4. Reducing time between runs. When an EDP installation reaches the point where many of its hour or longer runs are reduced to fifteen or thirty minutes, the percentage of the total time consumed by tape changing and other between run activities increases to a serious extent. Several things can be done to alleviate the problem: Aggressive methods work by operations, good scheduling, operations-oriented programing. Another solution is the use of duplicate tape drives, but this approach is rather expensive for most installations.

Cut down needless print-out

5. Combating the accumulation of data and inventory. The electronic computer's ability to produce data far exceeds the organization's ability to properly use it. Many people are still suspicious of magnetically stored information because they cannot see or feel the data. The tendency is to set retention cycles longer than necessary and to "hold" large numbers of tapes. This creates inventory problems, and the cost of keeping information on file defeats one of the purposes of magnetic tape. In the long run, the most effective solution to this problem is selling clients on the principle of management by exception.

6. Keeping current. Keeping up to date in this dynamic field can be a problem for an EDPM installation, but the problem is decreased when the changes are gradually and continually made. Improvements may bring new problems, however. The farther away from the actual machine language and the closer to the spoken-written language that programing comes, the more the programer's knowledge of what actually happens in the machine diminishes. But it is precisely this "what really happens" that is vital in truly effective programing. The manager must insure that his programers are provided with a sound foundation of computer and programing facts.

PRODUCTION PLANNING FOR COMPUTERS

J. W. Mitchell
THE ACCOUNTANT, April 2, 1960; pages 392-394

In planning for full utilization of a computer department, the manager should figure on a potential capacity of 168 hours a week. Routine maintenance will occupy from five to twenty-eight hours, leaving 140 hours available for routine operations. These can divide into three shifts of 44 hours weekly, leaving eight hours to spare. "If forty hours processing time is aimed at for each shift, a reasonable percentage of shift time is available for absorption of maintenance due to breakdown during processing. If this time is insufficient, the additional eight hours can be utilized in enabling the centre to catch up when operations are resumed."

It is assumed that programers will work only during the day shift, and access must be granted them for debugging purposes. This time should be specified to prevent them from monopolizing the computer whenever they wish. Night shifts should be restricted to routine processing on proved programs.

Computer department should charge for its work

For the purpose of figuring departmental costs, the department should be set up as an independent organization, charging other departments for its work. In this way, management may compare its charges with the expense of operation to ensure profitable operation, and the department managers may assess the benefits to their departments in reduced overhead expenses after meeting computer department charges.

Charges should be based on the equipment used and time occupied. Peripheral operations should be charged separately. Charges for work carried out in several departments, but more economically considered as a single computer run, should be shared equally by the departments. When equipment is shared for different jobs (for example, a sorting operation on the magnetic tape units while the computer is working on another job), the value of the work is the same, and should be charged as a full-time use of the equipment. Whatever saving accrues from such use of equipment, should belong to the computer department for its skill in programing and work planning. Where several programs always time-share the memory, "the hourly charge should be rated to take into account the expected number of programs against the equipment value."

In planning the computer usage, shift operation should be considered first, then maintenance time related to operation time. The staffing should be arranged so that the majority of the staff are on the day shift. This shift can be used for complicated and short-run programs, debugging, and data preparation. The other shifts should be used for routine processing and longer runs. Maintenance time can be scheduled around this framework. Adjustable wall charts showing equipment and staff available, work to be done, load planning for equipment, and other information should be provided.

Plan for debugging between runs

Machine time should be set up in advance, based on experience of the run, and allowing for seasonal fluctuations and breakdown. Debugging time can be scheduled in between the runs. Non-routine maintenance occupies only about two percent of computer time, however, occasionally a breakdown may take several hours to trace and rectify. "This entails considerable replanning of time, and, possibly, the usage of emergency routines or manual processing to get the work through on time. The emergency routines must be well planned to come into operation at a moment's notice, and should be periodically checked to ensure that they can be operated by the staff in the time. Recovery routines.... should also be well proven."

Information derived from log-books and time registers for the various items of equipment can be used in a time study of the most profitable use of the equipment. "It should be possible to arrive at a basic costing system for use with standard job-costing techniques in establishing day-to-day records of the equipment's use, together with a reliable basis for cost budgeting." The costing can be used for estimating future jobs and for establishing contract profitability.

COMPUTERS MUST KNOW

UNIVAC REVIEW, Winter 1960; pages 26, 27

The Chesapeake and Ohio Railway computing center has solved the problem of maintaining instruction manuals (covering electronic computer systems and allied routines) by installing Kardex files in each machine area. Instruction cards are inserted in the Kardex with the indexing or identifying information showing at the bottom edges. Each panel holds 69 cards. The panels are hung adjacent to the particular equipment at each work station, and contain numerically-sequenced cards covering each job done at that point. When an instruction is to be changed it is an easy matter to remove the card and replace it with the corrected one. Look-up by the operating staff is far easier in the Kardex than in a bound manual.

AN EVALUATION OF EQUIPMENT FOR EXPANSION OF DATA PROCESSING FACILITIES

DATA PROCESSING (U.S.A.), March 1960; pages 17-22

A set of requirements was devised by the U. S. Army Signal Supply Agency for redesigning their outgrown computing system and evaluating equipment manufacturers' proposals. The requirements were weighted according to their relative degree of importance. They were:

1.	Ability and capacity to perform data processing	25
2.	Recurring monthly cost	20

3.	Ease of programing	20
4.	Ease of conversion	15
5.	Ability to have expanded capacity in future	10
6.	Degree of back-up of equipment and exchange of information	10
7.	Miscellaneous company services	5

Information invitations were sent to eight leading manufacturers of business-type computers. Six companies accepted. Each company's systems group requested information from the Army as it needed it. This included descriptions of the basic daily computer job which was used as the "bench-mark" problem, supplemented by data sheets which provided transaction volumes, file sizes and formats, number of current average processing steps per input per run, number of instructions per run, printer requirements, computer room layout. The manufacturers' proposals were then evaluated according to the weighted requirements mentioned earlier. Details of the bench-mark problem were:

A bench-mark problem

1.	Prepare	1,000	Shipping Order/day (avg. 2 line items each)
	from	600	Requisitions/day (avg. 4 line items each)
	drawn on	195,000	Item Master Records
	and	44,000	Due in/out Records.
2.	Post	19,000	Depot Stock Accounting Transactions/Day
	against	195,000	Item Master Records.
3.	Post	9,000	Miscellaneous File Maintenance Transactions/day
	against	195,000	Item Records
	and	44,000	Due in/out Records.
4.	Answer	1,000	Stock Status Inquiries/day
	against	195,000	Stock Items Records
	and	44,000	Due in/out Records.
5.	Maintain	20,000	Customer Requisition Header Records (Name, Address, etc., required for printing shipping orders).
6.	Perform supply position study on	8,000	Active Stock Items/Day.
7.	Produce	400	Supply Action Recommendations/day.
8.	Produce	1,000	Availability Statements/day.
9.	Produce	2,500	Requisition Action Reports/day.

MANAGEMENT RAISES ITS ADP SIGHTS

AUTOMATIC DATA PROCESSING SERVICE NEWSLETTER, April 18, 1960

Advanced managements are realizing the need for a new way of thinking about information. This is the evolution from "management by exception" into "management by objectives," with its accent on decision-making rather than evaluation of operational results. This emphasis on intelligent decision-making is affecting the planning and installing of computing systems by focusing attention on what information is of real worth, and by centralizing and giving higher corporate status to the whole information responsibility. At present, the idea is so new that there is no agreement on where the information function should report in the organization chart. At International Latex Corporation, the group headed by Lionel Griffith which is setting up the information system, was told "to visualize the totality of the business, and then to formulate information handling for all decisions and control. This has resulted in... a three-dimensional model of rational information flow... a highly centralized data processing 'hub.'" Under this system, a distinction is made between the "intellectual" content of work in any area of activity, and the "mechanical" methods of providing the information for the intellectual operation. Plans also call for a "war-room" where all intelligence is centralized, as well as a novel use of sound, wherein basic management data, continually up-dated, will be available on sound tracks. Executives, instead of requesting a report, may dial into the system and receive the latest intelligence over the phone.

PREPARATION BEFORE AUTOMATION

UNITED STATES INVESTOR, April 11, 1960; pages 43, 44, 53

PREPARING FOR AUTOMATION

BANKING, April 1960; pages 45, 118

These two articles are reports of a talk given by Dr. F. Byers Miller, executive director of NABAC. Questions which a bank should consider in automating its savings or mortgage operations are:

1. Can customer service be maintained at its current speed and present level of quality?
2. Are all accounts numbered--and if not, what will it cost?
3. How much will it cost to code and qualify uncontrolled items?
4. In a commercial bank, what is the compatibility of the proposed system and equipment for an integrated banking system?
5. How can such equipment be used to process large volumes of original media for loan payments?
6. What are the cost and time factors in transferring balances on existing accounts to new records?

7. What forms will be needed and how much will they cost?
8. What are the physical site requirements?
9. What provisions must be made for displaced personnel?
10. What provisions must be made for emergency service in case of equipment failure?
11. How can gross errors in projected processing time be avoided?
12. How much will it cost to operate the equipment, taking into account idle time?
13. What will it cost to run in parallel during installation of the system?
14. What peripheral equipment will be needed?

CHECKING ACCOUNT AUTOMATION TODAY

J. Lewis Nungesser, Philadelphia National Bank
UNITED STATES INVESTOR, April 11, 1960, pages 39, 40

The reasons that banks did not rush to install the MICR plan upon recommendation of the ABA, are given. These include:

1. The ABA Bank Management Publication Number 147 advised banks not to act hastily.
2. The MIC specifications were written for printing technicians, and were difficult for bankers to interpret.
3. Some of the banks which did contact printers found them either ignorant of the whole process, or skeptical of their ability to produce satisfactory results.
4. At the NABAC meeting late in 1959, a spokesman for the printing industry pleaded for more time and relaxation of some of the more stringent tolerances, which was done.
5. At the time of the adoption of the MICR specifications, there was a lack of equipment to process the documents.

Difficulties in imprinting

As banks started magnetic ink imprinting either in their own print shops or through commercial printers, they ran into a number of difficulties. For example: one bank had a 30% reject rate on its first group of magnetically encoded documents. Banks found encoding deposit tickets, particularly in multiple-sets, a difficult and expensive job. Most banks were chagrined to find the large degree of inaccuracy in their files which was uncovered when customers were contacted to verify account titles. The transaction code posed some unexpected problems, as did the check-digit required in the MICR code. In addition, the public had to be educated to accept the new check design, although this has not been a difficult problem, on the whole.

Solutions are being found to these problems, and banks are reducing the reject rate to as low as one percent. The Federal Reserve System has encouraged magnetic ink encoding, and every member bank has been officially requested to magnetically encode its ABA number and routing symbol on all checks.

There are reactions to the MICR system. One large manufacturer does not offer magnetic ink character sorter/readers. A dual-code system has been offered by one equipment manufacturer, combining MICR for bank identification with perforated holes for account number and amount. Two banks have announced that they will use equipment in which paper documents are enclosed in plastic carriers for internal processing, eliminating the MICR feature. Claims have been made, too, that new and less expensive magnetic ink devices and devices that can optically scan conventional ink printing are available.

Finally, the answer for smaller banks appears to be the data processing centers which are appearing. IBM, RCA, and NCR all have announced plans for service bureaus capable of servicing banks. And at least two banks have made plans to perform check deposit processing for their correspondent banks.

WHICH CODE WILL BE USED FOR BRITISH CHEQUES?

OFFICE MAGAZINE, April 1960; pages 346, 347

British banks must soon make a decision as to which printed code they will use for electronic hangling of checks. The Committee of the London Clearing Bankers Electronics Sub-Committee has invited the National Physical Laboratory to do independent tests on three codes: the American Bankers Association code, the E. M. I. Ltd. figure-reading electronic device (FRED) code, and the C. M. B. code of Bull (France).

BANKING OPERATIONS IN THE COMING AGE OF AUTOMATION

BANKING OPERATIONS OF THE FUTURE

Neal J. Dean, Booz, Allen & Hamilton, New York

These two reports continue the author's suggestions on the future aspects of banking, including a discussion of the present status and relative merits of on-line and off-line data processing, the possibilities of the banking industry becoming in reality a financial utility, savings accounts assuming the functions of the present checking accounts for the average individual depositor, small banks going into automation via the cooperative computing center, and the future of banking 20 to 30 years hence. These are interesting reports, of value not only to the banking industry, but to all who use banking services. For copies of the reports, write to the author, Booz, Allen & Hamilton, 135 South La Salle Street, Chicago 3, Illinois.

NUMERIC VS ALPHABETIC CODING OF ACCOUNTS

John R. Moyer, NABAC Research Institute, Chicago
AUDITGRAM, April 1960; pages 4-7

The article discusses the three most widely used coding systems: straight numerics; straight numerics by account groups, and alphabetic sequenced numerics. In the first system, straight numerics, each account is assigned a number of predetermined length without regard to the alphabetic location of the account, its activity, or classification. With this system, cross-indexes are needed to provide quick access to all accounts. In a computer system, the computer must match a transaction with an account before updating the account. This means that the computer must look at each account number, including those which are inactive, a practice which wastes computer time, and results in increased wear on tapes and machines.

Numerics by account groups has advantages

The straight numerics by account groups alleviates this problem. The accounts are classified into major classifications, then into heavy or light activity, special handling, weekly or monthly statement preparation, etc., and numbered accordingly. Four additional digits would then provide identification of the account within the assigned classifications. Each account classification would be placed on its own magnetic tape file, with a resulting saving in sorting and merging time. Other advantages include: 1) corporate payrolls can be posted and reconciled simultaneously, 2) item sorting techniques for bookkeepers under any posting system are improved, and 3) time and costs are reduced on peripheral equipment.

The alphabetic sequenced numeric system has built-in difficulties. Numbers are assigned to accounts in alphabetic sequence, without regard to type of account. Here the bank must know the total number of accounts to be coded, the estimated net increase in the number of accounts each year, the gap spread between numbers, and the probability or odds to be accepted that x years from now there will still be a sufficient gap between assigned numbers so that the alphabetic sequence will be preserved.

Alphabetic sequencing has inherent problems

A description of the "classical occupancy" formula is given whereby a bank can figure the chances of having more than x number of accounts for any one gap in the number series. "The alphabetic sequence numeric account coding system is particularly suited to manual and semi-automated procedures. It is not so well suited for use in electronic data processing systems, although alphabetic grouping may be of value in preparing lists for other banking purposes. However, the installation of alphabetic sequenced coding is an excellent interim step towards full electronic processing since it provides for easy conversion of present systems to straight numerics. There is a very good reason why alphabetic sequence coding is not suitable for computer use: The extra digits necessary to identify corporate, individual, or heavy activity accounts could be costly in terms of machine time."

Each of the systems has many variations, and all must be analyzed in relation to the types of equipment available, particularly in relation to: 1) the type of input or data supply to the data processing system, 2) the type of output for convenient use by bank personnel, and 3) the internal processing advantages and disadvantages of the various systems.

"NAME CODE"

B. W. Taunton, First National Bank of Boston
DATA PROCESSING (U.S.A.), March 1960; pages 23, 24

In stock transfer work every transaction must be referred to the file and a number assigned to it. This is costly and time-consuming. The First National Bank of Boston has designed a method of coding accounts so that the computer can generate its own key for each account. The code serves to maintain the file in alphabetical order and is sufficiently unique for the machine to locate accounts in the file, post to them, and insert new accounts in their proper position.

Typist codes title during record preparation

"Name Code" is a set of rules under which the computer automatically selects 16 alphabetic characters from the stockholder's name, or the legal title in which the stock is registered. These characters are thereafter used by the computer for sorting, filing and posting operations in place of an ordinary account number. The basic work in the title, under which the item is to be filed, is keyed so the computer can recognize it by the typist's placing one or two colons in the proper location in the title when she prepares a new stock certificate or a record of transfer. As a by-product of this typing operation, punched paper tape is prepared as input to the computer system. Some examples of the way in which names are coded are included in this article.

The system has been operating for several months, and the bank has found it effective in its situation because of the relatively small volume of daily transactions--a little more than half of one per cent of the accounts. "The calculation of the key is relatively slow, by computer standards, and could become a significant factor in an active file. Hence, the code is not designed to substitute for numerical filing systems where they may be used. The latter are normally the most efficient in machine operations."

((See DPD: November 1959; page 12--"The First National Bank of Boston Has Novel Computer Operation."))

ENR SPECIAL REPORT ON COMPUTERS

ENGINEERING NEWS RECORD, April 14, 1960, pages 39-63

In a special report on computers in engineering office practices, it is stated that designers and construction organizations using computing facilities find they must modify their methods of collecting

and handling data, time schedules, techniques of problem solving and sometimes even their organization. Profitable use of computers requires restudy of existing practices and considerable advance planning.

When to consider using a computer

In the article, "Decisions: Buy? Rent? How to Organize for Profitable Operation," it is suggested that use of a computer is worth investigating when:

1. Accounting, mathematical, design or construction problems are so complex or such great accuracy is required that it would take too long for highly paid accountants or engineers to arrive at a solution.

2. So many solutions are possible for a specific problem that determination of the best solution by manual computation is not economically feasible.

3. The accounting or financial staff is spending too much time turning out repetitive reports.

4. Engineers are spending too much time on routine calculations.

5. The engineering staff is not large enough to take on added work and expansion is not desirable.

6. Large amounts of data are available but are useless because they are not correlated or tabulated.

As a rule of thumb, it is suggested that if an office has more than 20 engineers, it probably can support a computer of its own. But convenience and flexibility of an on-site computer should be compared with the larger equipment and other facilities of a computer service organization. Use of a service organization might be feasible if neither close contact between engineer and problem solution nor speed of communication between the firm and the service organization are important factors.

Buying may or may not be economical

If the firm decides to acquire its own machine, should it buy or rent? If adequate working capital is available, buying may be more economical. However these factors should also be considered:

1. Cost of renting is fully deductible from income tax.
2. A purchaser bears the cost of obsolescence.
3. It may be possible to rent first, buy later.
4. Rental cost includes a charge for maintenance and repair.

MONTHLY COST

	%
Computer (60-month amortization)	65
Maintenance (service only)	13
Parts	3
Paper, tape, typewriter ribbons	2
Power	2
Office space	2
Taxes, Insurance	13
	100

The computer organization may be of the closed shop or open shop type, or a combination of the two. The simpler the programing, coding, and operation of a computer are, the greater are the potentials for an open shop.

Other articles in the special report are: "Second Generation of Computers: Smarter and More Sophisticated," "How to Program a Digital Computer," and "How to Select the Right Computer."

FORMS CONSIDERATIONS IN SDA

NAVY MANAGEMENT REVIEW, March 1960; pages 8-10

The prerequisites of forms design for a Source Data Automation system involve 1) a detailed study and analysis of the circumstances in which the system is to operate, and 2) a knowledge of the features of the equipment that will be used to automate the system. Certain basic design principles that must be adhered to are:

1. Horizontal and vertical space requirements and limitations of the equipment must be considered.

2. Arrangement of data should be from left to right and from top to bottom, and tabular stops must be arranged conveniently.

3. Sequence of data and their relative position on all forms must be compatible with each other and with the writing equipment used.

4. Grouping of data for speed and accuracy is based on such facts as: constant data enters automatically; variable data manually.

5. Forms construction influences the efficiency of the writing operation and the use of the various copies. For example, time studies have proved that continuous forms are more efficient than other types of forms.

ELECTRONIC DATA PROCESSING—SUBJECT BIBLIOGRAPHY OF PERIODICAL LITERATURE—1959

Published by Lybrand, Ross Bros. & Montgomery

This bibliography is composed of referenced articles from seventy-five periodicals published during 1959. The articles are classified according to subject, and listed within each classification alphabetically. Where several subjects were covered in an article, its title is repeated under each pertinent heading. There are 62 classifications, including (at random) Accounts Payable, Auditing and EDP, Banking, Experience with EDP, Insurance, Management Control, Programming, Sales Analysis, Transportation.

The article titles, authors, periodicals, dates and pages are given. They are not annotated or digested. A list of the periodicals is given at the end of the bibliography to assist the reader in obtaining reprints. The bibliography may be requested from Lybrand, Ross Bros. & Montgomery, Management Services Research & Consulting Division, 2 Broadway, New York 4, N. Y.

"COMPUTERCADE"

The University of Arizona has mounted some IBM equipment in a van, and is sending it around the state to high schools. The purpose of the program is to show Arizona high school students the importance of mathematics in daily life and how study of the subject can be important to their future careers. Further information on the program can be obtained from Jim Allen, Press Bureau, the University of Arizona, Tucson, Arizona. ((From a news release.))

AUTOMATION OF ASTIA

This report, published by the Armed Services Technical Information Agency (ASTIA), describes the three-step installation of a computing system to record and retrieve information about technical literature made available by ASTIA to subscribing agencies. The reason for the need to install the system is described in the Foreword as follows:

"The task of transition is an enormous one. It involves one of the largest collections of scientific and technical reports in the free world; reports that reflect the results of the major part of the United States Government research and development program during and since World War II. The task is complicated by the continuing receipt of new reports from current programs arriving at the annual rate of about 30,000 titles, and by receipt of almost 2,000 separate requests for reports from our holdings <u>every</u>

working day. Add to this the fact that there was little precedence for such a venture, no proven system to follow, almost every step a new one to be studied and carefully planned, and the enormity of the task becomes apparent.

It is the long range objective to eventually, through automation, provide a comprehensive bibliographic information and announcement service in such a manner that the scientist or engineer can have at his fingertips, at any given time, information in the ASTIA collections pertinent to his needs. Also envisioned is the automation of essentially every step in filling a request for a report, including its automatic reproduction. This report relates the beginning of ASTIA's effort in these directions."

The 50-page report describes the solution to the retrieval problem. It may be ordered from the Office of Technical Services, U.S. Department of Commerce, Washington 25, D. C. Price: $1.25.

Systems Design

AUTOMATIC DATA PROCESSING AND THE OFFICE OF NAVAL MATERIAL

LCDR Archie B. Meihls, USN
MONTHLY NEWSLETTER, NAVY SUPPLY CORPS, December 1959; pages 32-38

The world-wide Navy inventory management system is being kept up to date as new equipment and new systems become available. The entire system is being evolved as an integrated system. For example, "with-in the materiel management area, inventory management, production control, and procurement are interrelated with each other as well as with the financial management area.... The same basic data which the production function uses to obtain materiel are also employed by the inventory manager as a basis for procurement, either from system stocks or through purchase action. These basic data are also employed by fiscal personnel to commit, obligate and expend funds, as well as to prepare cost, appropriation and other reports." A chart shows how various functions overlap so that, for example, "both inventory and fiscal records could easily be updated concurrently by a single computer....[and] materiel management is closely related to financial management in terms of automatic data processing systems...." In addition, although "the personnel area is not as directly related to materiel management...it has many elements of data common to financial management."

Unfortunately, there has not been any deliberate planning within the Navy for integration of these functions. "Whether this is good or bad cannot be fully determined, but certainly the unplanned and uncoordinated development of a production line by the various production foremen would not be tolerated by an effective industrial

organization. . . . The role of the Office of Naval Material and of the other functional managers in the Navy is therefore quite clear. . . . begin planning now for the future development of a coordinated management information system."

Applications

COMPUTING SYSTEM FOR MARINE CORPS PERSONNEL

Three NCR 304 computing systems have been set up across country to handle Marine Corps personnel facilities. The system figures the number and type of Marines needed to bring each unit up to combat strength. The system processes more than one million changes annually in personnel records, helps prepare training plans, budget and allotment checks, and other personnel record-keeping tasks. The system also can locate Marines with special skills for specific tasks. Should one or two of the three widely scattered systems be unable to function, the remaining system can take over the entire record-keeping job. ((From a news release.))

CARNEGIE TECH COMPUTING CENTER

Carnegie Institute of Technology will install a Bendix G-20 data processing system in the fall. "One of the unique aspects of the G-20 is that it can be remotely controlled from different buildings on the campus. Thus, the various research centers can supply data to the computer and evaluate results from operating positions outside the computation center" (Dr. Perlis, head of the Computation Center). ((From a news release.))

THE BIG COUNT IS UNDER WAY

MANAGEMENT AND BUSINESS AUTOMATION, April 1960; pages 16-19, 40

The processing of the census sheets is now underway. The large sheets worked on by the enumerators are microfilmed in the Census Bureau's Operations Office in Jeffersonville, Indiana. These are reduced to one square inch on reels of microfilm and are then transported to the FOSDIC installation at Suitland, Maryland. The photoelectric scanners of FOSDIC will read the data recorded on the original sheets as blacked-in circles. FOSDIC is able to distinguish between an intentional mark and an erasure, and in the case of duplicate markings in the same grid area, selects the denser of the two. The information is recorded on magnetic tape to be processed by the Bureau's two Univac 1105 computers and two additional 1105's at the

Illinois Institute of Technology and the University of North Carolina. Univac system high speed printers will convert output tapes into printed reports which are expected to total about 100,000 pages. It is estimated that the reports will be completed six to 18 months sooner than the 1950 census, and will save about $15 million over the previous reports.

Programing

HONEYWELL FACT SYSTEM

FACT (Fully Automatic Compiling Technique) is the Honeywell 800 automatic programing system. "FACT translates problem-oriented language into compiler language", creating the operating program, called the object program, in the form of punched cards or magnetic tape. In addition to the object program, FACT produces the following outputs: 1) a listing of the narrative source statements; 2) a detailed map, showing the use of the high-speed memory by the object program; 3) diagnostic comments describing the number, kinds, and locations of all errors encountered in the source program (such as incorrect spelling, improper use of data names, illegal file structure); 4) operator instructions requesting specific actions during and immediately after compilation; 5) new file directories.

A 94-page manual, titled "FACT--a New Business Language," may be obtained from Minneapolis-Honeywell, Datamatic Division, Wellesley Hills 81, Mass.

ROOTING OUT DATA THAT DISTORT BUSINESS FORECASTING

BUSINESS WEEK, March 19, 1960; pages 190-194

A new computer program has been developed by Julius Shiskin at the Census Bureau for "the decomposition of a time series into seasonal, cyclical, and irregular components." The new economic analysis routine is being made available by four computing service organizations for private business use.

Equipment

MODEL 160 MAKES DEBUT

Control Data Corporation, Minneapolis, has announced the availability of Model 160, a small computer in the $60,000 price class. Standard input-output equipment consists of paper tape reader and paper tape punch. Optional equipment includes an electric typewriter, magnetic tape units (15,000 or 30,000 characters per second), card reader and punch, and line printer. The computer has a 1496, twelve-bit word core memory and a basic add time of 12.8 microseconds. ((From a news release.))

BENDIX G-20 COMPUTER

The Bendix G-20 computer system was designed to solve large scientific and data processing problems with comparatively inexpensive equipment. The computer can "delegate authority" to subordinates called control buffers which proceed on their own, reading punched tape or cards, looking up data on magnetic tape and printing answers. Meanwhile the central processor can work on other portions of the problem or instruct other subordinates. The control buffer signals the central processor when it has completed its assigned task. If the central processor is not ready for the information, or is working on another task, it retains the signal and turns to the subordinate unit when it is ready to use it. The G-20 is modular and can be assembled to the size desired. The control buffers may be located close to the central processor or at a geographically remote location.

The normal input-output devices for card and paper tape and 60,000 character per second magnetic tape units are available. Basic add time is 7 microseconds. ((From a news release.))

THE FERRANTI ORION DATA PROCESSING SYSTEM

O & M BULLETIN, February 1960, pages 22, 23

The Orion Data Processing System is a transistorized computer with a magnetic core storage of up to 16,384 words. Additional storage may be provided by any number of magnetic drums, each holding the same number of words as the main store. A word is 48 bits plus a parity bit. The Orion code has more than 100 functions, and each instruction is single-word length with a 3-address or modified 2-address. The system has a time-sharing facility which allows several independent programs to be run concurrently. Each of the peripheral units is connected to a single-word buffer store and can read or write into the magnetic core store directly. The price range is Ł100,000 to Ł300,000, depending on the peripheral equipment used.

Comment

NOTES ON DATA RETRIEVAL

Data retrieval is conceptually simple: a large file of reports or documents contains information about some particular group of subjects. People then wish to study certain subjects with the assistance of pertinent reports from this file. The data retrieval problem is: how to insert the reports into the file and call them out so that the users have a high probability of getting all of the reports pertinent to their questions and a low probability of getting those which do not apply.

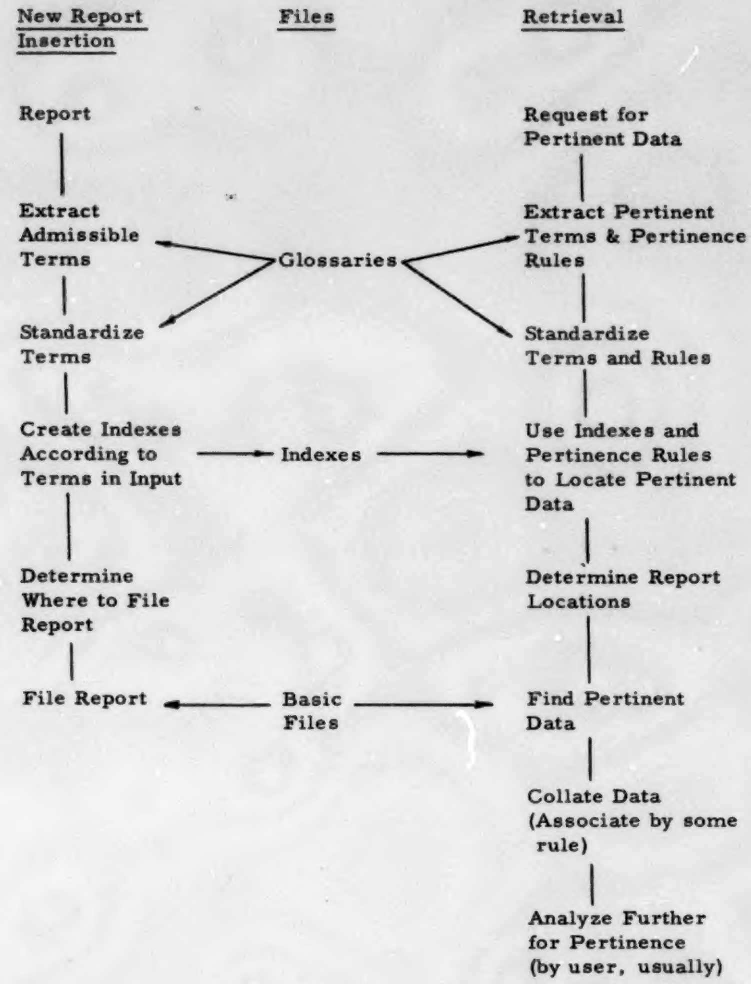
One solution is a straight-forward method used in most business file maintenance processes: Whenever a request is made, search all reports in the file, extracting those which are pertinent. If the user (or previous process) can define one key characteristic of the reports, such as a part number or payroll number, the process can be made more economical. Files can be organized in sequence by this key number to make searching for a given report faster. However, if the requester cannot define the key, a search of all records must be made. An example is the request for all personnel with skills as a machinist, or for all parts which have been out-of-stock over one week.

How to find what you can't define

Experts refer to "data retrieval" as those situations in which the users cannot define a convenient key word and in which the file of reports is very large and so cannot be searched exhaustively in reasonable times. Examples are: the retrieval of technical information , patents, military intelligence data, court findings in law, etc.

Basic to all of these areas is the need to devise methods (and equipment) to permit the location of pertinent data rapidly and economically, in particular, a method which searches through as few non-pertinent reports as possible before finding the pertinent ones.

Data retrieval problems have some common characteristics, diagramed below, which may be used as a checklist, in analyzing the statement of data retrieval problems and proposed solutions. The two basic procedures are the insertion of a new report, and the retrieval upon request:



Terms. Using technical data in the EDP field as an example, typical terms might be: "computer," "file," "operator," "1957." Extraction and standardization of terms (and rules) is often a manual step even in automatic systems, since it is difficult, for example, to program a machine to interpret "a stack of reports" as "file."

Apply the pertinence rule to find a "set" of information

Pertinence Rule. A typical pertinence rule might be: find all documents pertinent to "computers with tapes used for inventory control since 1957." Internally, the system might denote this rule as: find reports with these terms (related by the underlined words): "computer" and "tapes" and ("inventory control" or "production control") and (1957 or 1958 or 1959 or 1960). "Production control" is a term determined by the system to be sometimes synonymous with, or to include "inventory control" and therefore also pertinent. The complexities of useful indexes can be appreciated.

Data Collation. Collation of data is often merely the sorting or permutation of key terms, followed by sorting. Often the basic file is in two parts, a concise standardized summary of the report and the complete input as received. This adds to the complexity of indexes and the put-away and actual data retrieval steps; but reduces the time to search in the summaries.

Updating is a required characteristic of a file

Two other procedures occur. First, it must be possible to update the system. It is characteristic of a store of data that the user is continually reorganizing and adding to it, in his mind, and wants the retrieval system to change correspondingly. Thus it is necessary to add new terms, new pertinence rules, new collation rules; revise indexes, if they imply an organization of the data, etc.

Second, especially in military applications, it is necessary to find internal relationships between the reports. This might be thought of as taking one report at random and trying to retrieve and collate all other reports which relate to the first by some rules of pertinence (same area, same activity, temporal relations, etc.)--a variation of the normal retrieval procedure.

The object of most studies in data retrieval at the present time is to devise index methods to permit finding data related to a given input which consists of several terms and sometimes pertinence rules. Generally an automatic retrieval system can be developed, if the need is pressing enough (e.g., military); but often solutions which make use of presently available equipment are very expensive. This problem is under constant study and should be watched by those who use data from large data stores or files.

We understand a well-known breakfast food company is going into the computer manufacturing business. We assume it will be a CEREAL machine.

Training

Operations Research Methods

Date:	June 6-16, 1960
Place:	Purdue University
Fee:	$200
Information:	University Extension Administration, Purdue University, Lafayette, Indiana

Work Design, an experimental course sponsored by St. Louis Chapters of AIIE and SAM, and Washington University

Date:	June 6-17, 1960
Place:	Washington University, St. Louis, Missouri
Course requirement:	Only those with previous training in some procedure of studying work should attend.
Fee:	$175
Information:	University College, Washington University, St. Louis 30, Mo.

University of Michigan Summer Courses, 1960, a program of intensive non-credit courses for practicing engineers and scientists including electronic data processing

Date: June 13-24, 1960

Place: Ann Arbor, Michigan

Information: R. E. Carroll, Coordinator, Engineering Summer Courses, 2038 East Engineering Building, The University of Michigan, Ann Arbor, Michigan

Cornell University Industrial Engineering Seminars

Date: June 14-17, 1960

Place: Cornell University, Ithaca, New York

Subjects: Industrial Management, Engineering Administration, Operations Management of the Smaller Company, Work Measurement, Systems Simulation Using Digital Computers, Statistical Decision-Making; Theory and Applications, Statistical Reliability Analysis: Theory and Applications

Information: J. W. Gavett, Seminars Coordinator, Upson Hall, Cornell University, Ithaca, N. Y.

EDP Installations in Operation--A Conducted Tour of Practical Applications, sponsored by San Diego State College

Date: August 8-20, 1960

Place: Tour starts at San Diego State College and includes Los Angeles, and vicinity, San Francisco, and vicinity, with sightseeing stops at Sequoia and Yosemite.

Content: Visits to large and small computer installations and computer manufacturers. Persons without automation background can benefit, as well as those who are familiar with the subject.

Fee: $135 plus meals (including bus fare and accommodations)

Course requirements: Credit for Upper or Graduate Division Status. For audit: college matriculation and consent of instructors

Registration: Dr. E. Dana Gibson, Professor, Office Management, San Diego State College, San Diego 15, California

American Management Association Seminars--Office Management Section

Date: August 15-19--"Data Processing Function"
"Information & Reporting Systems"
August 29-31--"Data Processing Evaluation"
And others in the systems and procedures field

Information: American Management Association, Seminar Registration, 1515 Broadway, New York 36, New York

A Development Program in O. R., sponsored by Case Institute of Technology

Date: September 20, 1960 to January 27, 1961

Place: Case Institute of Technology, Cleveland, Ohio

Information: Dr. E. Leonard Arnoff, Asst. Dir., Operations Research Group, Dept. of Mgmt., Case Institute of Technology, University Circle, Cleveland 6, Ohio

Meetings

International Conference of Institute of Internal Auditors

Date: June 12-15, 1960
Place: Milwaukee, Wisconsin (Hotel Schroeder)
Information: Institute of Internal Auditors, 120 Wall Street, New York 5, New York

International Accounting Conference, sponsored by National Association of Accountants

Date: June 19-22, 1960
Place: New Orleans, La. (Hotel Roosevelt)
Information: National Association of Accountants, 505 Park Avenue, New York 22, New York

National Machine Accountants Association National Conference

Date: June 22-24, 1960
Place: San Francisco, California

On Tuesday, June 21, a pre-conference tour will be conducted to outstanding data processing centers in nearby areas. In addition to the regular meetings, a Hall of Discussions will be held during the conference which will feature questions of general interest discussed informally. Following the conference, registrants may continue to Honolulu for a two-day meeting beginning on Monday, June 27, featuring data processing experts from the Pacific Basin and Asia.

Information: Address inquiries to 1960 National Conference, NMAA, P.O. Box 3617, Rincon Annex, San Francisco 19, California

"Automation in Business Decision Processes," sponsored by Los Angeles Chapter of ACM

Date: June 23, 1960
Place: University of California at Los Angeles (UCLA)
Information: Owen R. Mock, North American Aviation, Inc., Los Angeles 45, California

Conference, British Computer Society, Ltd.

Date: July 4-7, 1960
Place: Harrogate, Yorks, England
Information: Miss D. E. Pilling, Electronic Computing Laboratory, The University, Leeds 2, England

Symposium on Computers and Data Processing, sponsored by University of Denver

Date: July 28, 29, 1960
Place: Estes Park, Colorado (Stanley Hotel)
Information: W. H. Eichelberger, Denver Research Institute University Park, Denver 10, Colorado

Bendix G-15 Users' Exchange Organization Annual Conference

Date: August 10-12, 1960
Place: Pittsburgh, Pennsylvania (Pittsburgh Hilton)
Theme: "Strategic Programming"

National ACM Conference
- Date: August 23-25, 1960
- Place: Marquette University, Milwaukee, Wisconsin

SHARE XV Meeting
- Date: September 12-16, 1960
- Place: Pittsburgh, Pa. (Pittsburgh Hilton Hotel)
- Information: E. B. Weinberger, Gulf Research & Development Co., Drawer 2038, Pittsburgh 30, Pa.

Univac Users Association
- Date: September 22, 23, 1960
- Place: Washington, D. C.

CUE, Burroughs 220 Users' Group
- Date: October 4-6, 1960
- Place: Philadelphia, Pennsylvania

Electronic Computer Exhibition and Business Symposium
- Date: October 4-12, 1960
- Place: London, England (Olympia)
- Information: Mr. D. C. Scoones, Peat, Marwick, Mitchell & Co., 94-98 Petty France, London SW 1, England

NABAC National Convention
- Date: October 10-12, 1960
- Place: Los Angeles, California
- Information: NABAC, The Association for Bank Audit, Control and Operation, 38 South Dearborn St., Chicago 3, Illinois

International Systems Meeting, sponsored by Systems and Procedures Association
- Date: October 10-12, 1960
- Place: New York, N. Y. (Hotel Commodore)
- Information: Systems and Procedures Association, 4463 Penobscot Bldg., Detroit 26, Michigan

The Institute of Management Sciences (TIMS) International Meeting
- Date: October 20-22, 1960
- Place: New York City (Hotel Roosevelt)
- Subjects: Behavioral Science and Management Science, Applications and Tools of Management Science, Use of Computers in Simulation
- Information: Mr. James Townsend, 30 East 42nd Street, New York 17, New York

Computer Applications Symposium, sponsored by Armour Research Foundation
- Date: October 26, 27, 1960
- Place: Chicago Illinois (Morrison Hotel)
- Information: Andrew Ungar, Armour Research Foundation, 10 West 35th Street, Chicago 16, Illinois

References

DATA PROCESSING DIGEST does not provide copies of the original material digested or reviewed in this issue. The publishers addresses are listed below for your convenience in writing to them for more complete information.

The Accountant
4 Drapers' Gardens
Throgmorton Avenue
London EC 2, England

Auditgram
38 South Dearborn Street
Chicago 3, Illinois

Automatic Data Processing Service Newsletter
40 Wall Street
New York 5, New York

Automation
Penton Building
Cleveland 13, Ohio

Banking
12 East 36th Street
New York 16, New York

Business Week
330 West 42nd Street
New York 36, New York

Data Processing
956 Maccabees Building
Detroit 2, Michigan

Engineering News Record
330 West 42nd Street
New York 36, New York

Journal of Machine Accounting
1700 West Central Road
Mt. Prospect, Illinois

Management & Business Automation
600 West Jackson Blvd.
Chicago 6, Illinois

N. A. A. Bulletin
505 Park Avenue
New York 22, New York

Navy Management Review
Supt. of Documents
U. S. Government Printing Office
Washington 25, D. C.

Newsletter, Navy Supply Corps
Bureau of Supplies & Accounts
(N1), Room 0211 Main Navy
Washington 25, D. C.

Office Magazine
319 High Holborn Street
London WC1, England

O & M Bulletin
Treasury Chambers
Great George Street
London SW1, England

United States Investor
286 Congress Street
Boston 10, Mass.

Univac Review
Remington Rand Univac Div.
315 Fourth Avenue
New York 10, New York

DATA PROCESSING DIGEST is published each month by Canning, Sisson and Associates, Inc., 1140 South Robertson Boulevard, Los Angeles 35, California. Subscription rate: $24.00 per year. Foreign postage (exclusive of Canada and Mexico): $2.50 additional. Single copies: $3.00 when available. Executive Editors: Richard G. Canning and Roger L. Sisson. Managing Editor: Margaret Milligan.